

Sammanfattning av LV2 i Flervariabelanalys

Elon Andiné

William Karlsson
Marco Renzoni

Lovisa Petersson
David Sjöberg

Marcus Petersson

28 januari 2022

1 Clairauts sats

I början av den första lektionen under läsvecka 2 gick vi igenom Clairauts sats (Sats 9. under avsnitt 2.5. i Persson/Böiers: Analys i flera variabler) vilket kan ses nedan.

Sats: Om $f : \mathbf{R}^2 \rightarrow \mathbf{R}$, $f \in C^2$ i en omgivning till (a, b) , så är $f''_{xy}(a, b) = f''_{yx}(a, b)$.

Beviset till satsen kan sammanfattas genom att anta en kvadrat ABCD med punkterna $A = (a, b)$, $B = (a + h, b)$, $C = (a + h, b + h)$, $D = (a, b + h)$. Vi antar att h är så litet att kvadraten ryms i omgivningen där f är C^2 . Genom att ta fram funktionerna (1) och (2),

$$F(x) = f(x, b + h) - f(x, b), \quad (1)$$

$$G(y) = f(a + h, y) - f(a, y), \quad (2)$$

kan vi utav dessa utforma ekvationen (3),

$$F(a + h) - F(a) = G(b + h) - G(b). \quad (3)$$

Genom att utnyttja medelvärdsatsen i två steg för (3) och förenkling, kan ekvationen skrivas som

$$f''_{xy}(\xi, \zeta) = f''_{yx}(\theta, \eta), \quad (4)$$

där då $h \rightarrow 0$ så går $\xi \rightarrow a$, $\theta \rightarrow a$, $\zeta \rightarrow b$ och $\eta \rightarrow b$. Vilket insatt i (4) bevisar satsen.

2 Taylorutvecklingar i fler variabler

För en funktion $f : \mathbf{R}^2 \rightarrow \mathbf{R}$ där $f(x, y)$ ska utvecklas kring baspunkten (a, b) för $f \in C^{k+1}$, kan en envariabelfunktion, $F(t)$, bildas genom att fixera två konstanter, h_1 och h_2 , vilket blir

$$F(t) = f(a + th_1, b + th_2). \quad (5)$$

Nu med hjälp av (5) kan en taylorutveckling för en variabel göras, där F är kring $t = 0$, vilket ser ut som

$$F(t) = \sum_{j=0}^k \frac{F^{(j)}(0)}{j!} t^j + \frac{F^{(k+1)}(\xi)}{(k+1)!} t^{k+1}, \quad (6)$$

där $\xi \in (0, t)$. För att översätta (6) från $F(t)$ till $f(a+h_1t, b+h_2t)$ utnyttjas kedjeregeln för att ta reda på vad $F'(t) = D(f(a+h_1t, b+h_2t))$ är, vilket blir $F'(t) = (h_1 \frac{\partial}{\partial x} + h_2 \frac{\partial}{\partial y})f(a+th_1, b+th_2)$. Vi kan även se att för j :te derivata får vi $F^{(j)}(t) = (h_1 \frac{\partial}{\partial x} + h_2 \frac{\partial}{\partial y})^j f(a+th_1, b+th_2)$. $F^{(j)}(0)$ är därmed $(h_1 \frac{\partial}{\partial x} + h_2 \frac{\partial}{\partial y})^j f(a, b)$. Eftersom $\frac{\partial}{\partial x} \cdot \frac{\partial}{\partial y}$ kommuterar enligt Clairauts sats så kan Binomialsatsen användas för att uttrycka $(h_1 \frac{\partial}{\partial x} + h_2 \frac{\partial}{\partial y})^j$, fast det går även att förenkla genom att sätta $\mathbf{a} = (a, b)$, $\mathbf{h} = (h_1, h_2)$ och $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$. Insättning av dessa i $F(1) = f(\mathbf{a} + \mathbf{h})$ ger

$$f(\mathbf{a} + \mathbf{h}) = \sum_{j=0}^k \frac{(\mathbf{h} \cdot \nabla)^j}{j!} f(\mathbf{a}) + \frac{(\mathbf{h} \cdot \nabla)^{(k+1)}}{(k+1)!} f(\vec{\xi}), \quad (7)$$

där $\vec{\xi}$ är på sträckan mellan $\mathbf{a}$ och $\mathbf{a} + \mathbf{h}$.

3 Stationära punkter och dess karaktär

Antag att vi har en funktion av två variabler $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, och en stationär punkt $\mathbf{a} = \begin{bmatrix} a \\ b \end{bmatrix}$. Innebörden av en stationär punkt är att samtliga första derivator vid punkten blir noll, alltså $\nabla f(\mathbf{a}) = 0$. Antagandet leder till att vi kan omformulera och förenkla Taylorutvecklingen för funktioner av två variabler,

$$f(a+h, b+k) - f(a, b) = \frac{1}{2}(Ah^2 + 2Bhk + Ck^2) + O(\|(h, k)\|^3), \quad (8)$$

$$A = f''_{xx}, \quad B = f''_{xy}, \quad C = f''_{yy}. \quad (9)$$

I ekvation (8) så har vi flyttat över $f(a, b)$ till vänsterledet och $\nabla f(\mathbf{a})$ har tagits bort. Storleken av första termen i vänsterledet kommer nu att avgöra om punkten är en maximipunkt, minimipunkt eller något annat. Om $f(a+h, b+k) < f(a, b)$ så får vi ett maximivärde vid $\mathbf{a}$, eftersom vi har förflyttat oss en godtyckligt liten bit från $\mathbf{a}$ och fått ett mindre värde. På samma sätt får vi en minimipunkt vid $\mathbf{a}$ om $f(a+h, b+k) > f(a, b)$. Både storleken och tecknet av vänsterledet kan bli svåra att bestämma då funktioner blir allt mer komplicerade. Därför studeras istället högerledet. Högerledets tecken kommer helt och hållet domineras av den så kallade kvadratiske formen, och ordo-termen försummas. Den kvadratiske formen skrivs som,

$$Q(h, k) = Ah^2 + 2Bhk + Ck^2. \quad (10)$$

Med hjälp av tecknet på Q är det möjligt att bestämma karaktären av stationära punkten $\mathbf{a}$ på följande sätt:

Q är positivt definit $\Rightarrow \mathbf{a}$ är ett strängt lokalt minimum,

Q är negativt definit $\Rightarrow \mathbf{a}$ är ett strängt lokalt maximum,

Q är indefinit $\Rightarrow \mathbf{a}$ är en sadelpunkt,

Q är semidefinit $\Rightarrow \mathbf{a}$ går inte att klassificera.

Exempel

$$f''_{xx} = 1, \quad f''_{xy} = 2, \quad f''_{yy} = 6,$$

$$\Rightarrow Q(h, k) = Ah^2 + 2Bhk + Ck^2 = h^2 + 4hk + 6k^2.$$

Kvadratkomplettering ger förenklingen,

$$h^2 + 4hk + 6k^2 = (h + 2k)^2 - 4k^2 + 6k^2 = (h + 2k)^2 + 2k^2.$$

Både $(h + 2k)^2$ och $2k^2$ förblir positiva oavsett vad h och k väljs till att vara. Därför är Q positivt definit och $\mathbf{a}$ är därmed en minimipunkt.

Notera att en enkel ändring $f''_{yy} = 3$, skulle förändra karaktären av punkten. Slutresultatet hade med ändringen blivit $(h + 2k)^2 - k^2$, vilket ger ett indefinit tecken på Q . Punkten $\mathbf{a}$ hade således blivit en sadelpunkt.

4 Matrisformulering av kvadratiska formen

Denna del kommer behandla hur karaktären på kvadratiska formen kan bestämmas och då implicit även hur karaktären på stationära punkter kan avgöras.

4.1 För två variabler

Låt

$$\mathbf{h} = \begin{bmatrix} h \\ k \end{bmatrix}, M = \begin{bmatrix} A & B \\ B & C \end{bmatrix},$$

där A , B och C är samma som i (5).

Då kan den kvadratiska formen skrivas om så att

$$Q(\mathbf{h}) = \mathbf{h}^T M \mathbf{h}. \quad (11)$$

Man kan då avgöra teckenkaraktären på Q genom att undersöka determinanten till M

$$\det(M) = AC - B^2. \quad (12)$$

Följande påståenden gäller:

- $\det(M) > 0$ & $A > 0 \Rightarrow Q$ är positivt definit,
- $\det(M) > 0$ & $A < 0 \Rightarrow Q$ är negativt definit,
- $\det(M) < 0 \Rightarrow Q$ är indefinit,
- $\det(M) = 0 \Rightarrow Q$ är semidefinit.

Detta gäller dock endast för $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, så följande delkapitel kommer behandla hur karaktären på stationära punkter kan bestämmas i fler variabler.

4.2 För n variabler

För n variabler gäller att om M är en symmetrisk $n \times n$ -matris och $\mathbf{h} \in \mathbb{R}^n$, är $Q(\mathbf{h}) = \mathbf{h}^T M \mathbf{h}$ nu $Q : \mathbb{R}^n \rightarrow \mathbb{R}$.

Eftersom M är både reell och symmetrisk, betyder det att den är diagonaliserbar. Den kan alltså skrivas om så

$$M = TDT^{-1}, \quad (13)$$

där T är en ortogonal matris (dvs $T^T = T^{-1}$) med M 's egenvektorer som kolonnvektorer och D är en diagonal matris med motsvarande egenvärden till M . Om vi gör variabelbytet

$$\mathbf{y} = T^{-1}\mathbf{h} \Leftrightarrow \mathbf{h} = T\mathbf{y}, \quad (14)$$

ger det tillsammans med (13)

$$Q(\mathbf{y}) = \mathbf{y}^T D \mathbf{y} = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} \begin{bmatrix} \lambda_1 y_1 \\ \vdots \\ \lambda_n y_n \end{bmatrix} = \sum_{k=1}^n \lambda_k y_k^2, \quad (15)$$

där $\mathbf{y} \neq \mathbf{0}$ (ty annars är T ej inverterbar) och λ_k är egenvärdena till M .

Utifrån (15) ser vi att egenvärdena kontrollerar teckenkaraktären på Q . Följande sats kan nu formuleras.

Sats:

- (i) Alla egenvärden till M är positiva. $\Leftrightarrow Q$ är positivt definit.
- (ii) Alla egenvärden till M är negativa. $\Leftrightarrow Q$ är negativt definit.
- (iii) M har både positiva och negativa egenvärden. $\Leftrightarrow Q$ är indefinit.
- (iv) M har något egenvärde $\lambda_i = 0$. $\Leftrightarrow Q$ är semidefinit.

Eftersom M tydligen är en viktig matris, har den fått ett namn. Den kallas *hessianen* till f

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}. \quad (16)$$

Denna matris kan då evalueras i den stationära punkt som vill undersökas. Man kan alltså med genom att hitta hessianens egenvärden avgöra karaktären av stationära punkter.

5 Vektorvärda funktioner

Antag att det finns en vektor $\mathbf{f}$ som går från $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^p$. Därefter delas denna vektorn in i dess olika element f_i där varje element går från $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$. Detta för att det nu erhålles en funktion vi har sett tidigare och man kan då på ett smidigare sett linjärisera med avseende på en funktion enligt följande,

$$f_i(\mathbf{a} + \mathbf{h}) = f_i(\mathbf{a}) + \nabla f_i(\mathbf{a}) \cdot \mathbf{h} + \|\mathbf{h}\| \rho(\mathbf{h}) \quad (17)$$

där $\rho \rightarrow 0$ då $\mathbf{h} \rightarrow \mathbf{0}$

Därefter samlas alla partiella derivator ihop i en $p \times n$ matris, som består av ett system av funktioner. Denna matrisen kallas Funktionalmatrisen och är även känd som Jacobimatrisen,

$$\mathbf{f}'(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1} & \cdots & \frac{\partial f_p}{\partial x_n} \end{bmatrix}. \quad (18)$$

Kommande uttryck är ekvivalent med valfri rad från funktionalmatrisen,

$$\nabla f_i(\mathbf{x}) \cdot \mathbf{h} = \begin{bmatrix} \frac{\partial f_i}{\partial x_1} & \cdots & \frac{\partial f_i}{\partial x_n} \end{bmatrix} \begin{bmatrix} h_1 \\ \vdots \\ h_n \end{bmatrix}. \quad (19)$$

En sammanställning av dessa rader av funktionern $\mathbf{f}$ ger oss en taylorapproximering av vektorvärda funktioner enligt följande

$$\mathbf{f}(\mathbf{a} + \mathbf{h}) = \mathbf{f}(\mathbf{a}) + \mathbf{f}'(\mathbf{a})\mathbf{h} + \|\mathbf{h}\|\boldsymbol{\rho}(\mathbf{h}) \quad (20)$$

där $\boldsymbol{\rho}(\mathbf{h}) \rightarrow \mathbf{0}$ då $\mathbf{h} \rightarrow \mathbf{0}$.

6 Kedjeregeln för vektorvärda funktioner

Låt $\bar{g} : \mathbb{R}^n \rightarrow \mathbb{R}^p$, $\bar{f} : \mathbb{R}^p \rightarrow \mathbb{R}^m$ vara två differentierbara funktioner. Där $\bar{a}$ är inre punkt till $\bar{g}$ och $\bar{g}(\bar{a})$ är inre punkt till $\bar{f}$. Låt

$$\bar{t} = \begin{bmatrix} t_1 \\ \vdots \\ t_n \end{bmatrix}, \bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_p \end{bmatrix}, \bar{g}(\bar{t}) = \begin{bmatrix} g_1(\bar{t}) \\ \vdots \\ g_p(\bar{t}) \end{bmatrix}, \bar{f}(\bar{x}) = \begin{bmatrix} f_1(\bar{x}) \\ \vdots \\ f_m(\bar{x}) \end{bmatrix}.$$

Vi tittar på ett av elementen av $\bar{f}$, $f_i : \mathbb{R}^p \rightarrow \mathbb{R}$. Då får vi en flervariabelfunktion som är skalärvärd istället för vektorvärd

$$(\bar{f} \circ \bar{g})_i(\bar{t}) = f_i(\bar{g}(\bar{t})).$$

Genom att sedan partiellt derivera med avseende på en variabel t_j ger kedjeregeln oss uttrycket

$$\frac{\partial}{\partial t_j} (\bar{f} \circ \bar{g})_i(\bar{t}) = \nabla f_i(\bar{g}(\bar{t})) \frac{\partial \bar{g}(\bar{t})}{\partial t_j} = A.$$

Utrycket kan istället skrivas med vektorer som matrismultipliceras

$$A = \underbrace{\begin{bmatrix} \frac{\partial f_i}{\partial x_1} & \cdots & \frac{\partial f_i}{\partial x_p} \end{bmatrix}}_{\text{Rad } i \text{ av } \bar{f}'(\bar{g}(\bar{t}))} \left\{ \begin{bmatrix} \frac{\partial g_1}{\partial t_j} \\ \vdots \\ \frac{\partial g_p}{\partial t_j} \end{bmatrix} \right\} \text{ Kolonn } j \text{ av } \bar{g}'(\bar{t}).$$

Vi ser här att de två vektorerna är rad eller kolonnvektorer av Funktionalmatrisen för $\bar{f} \circ \bar{g}$ respektive $\bar{g}$. Vi har endast tittat på ett element av $\bar{f}$, men med detta uttryck kan vi kombinera alla element för att skapa en formel för derivatan av $\bar{f} \circ \bar{g}$. Vi får då av matrismultiplikation uttrycket

$$(\bar{f} \circ \bar{g})'(\bar{t}) = \bar{f}'(\bar{g}(\bar{t}))\bar{g}'(\bar{t}).$$

7 Implicita funktionssatsen i n variabler

Låt $F : \mathbb{R}^n \rightarrow \mathbb{R}$ vara en C^1 -funktion och punkten $\bar{a} \in \mathbb{R}^n$ vara en punkt på nivå(hyper)kurvan $F(\bar{x}) = C$.
Om

$$\frac{\partial F}{\partial x_k} \neq 0 \quad (21)$$

så finns en öppen omgivning U av $\bar{a}$ så att restriktionen av nivåkurvan till U implicit definierar en C^1 -funktion $f_k : \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ som ger $x_k = f_k(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n)$.

7.1 Exempel

Med hjälp av implicita funktionssatsen kan vi till exempel visa att nivåkurvan i 2 variabler $F(x, y) = 1$ där $F(x, y) = e^y - x \cos(y)$ entydigt definierar en funktion $y = f(x)$ kring punkten $(0, 0)$.

För att $y = f(x)$ ska existera kring $(0, 0)$ måste punkten vara på nivåkurvan vilket kontrolleras genom insättning i $F(x, y)$

$$F(0, 0) = e^0 - 0 \cos(0) = 1. \quad (22)$$

villkoret

$$\frac{\partial F}{\partial y} \neq 0$$

kontrolleras också genom insättning

$$\frac{\partial F}{\partial y}(0, 0) = e^0 + 0 \sin(0) = 1 \neq 0$$

$\implies$ enligt implicita funktionssatsen att $y = f(x)$ existerar vid $(0, 0)$.